Development of Modified Force Field for Cation–Amino Acid Interactions: *Ab Initio*-Derived Empirical Correction Terms with Comments on Cation–π Interactions

OREOLA DONINI,[1] DONALD F. WEAVER[1,2]

[1] *Department of Chemistry, Queen's University, Kingston, Ontario K7L 3N6, Canada*
[2] *Department of Medicine, Queen's University, Kingston, Ontario K7L 3N6, Canada*

Received 13 November 1997; accepted 6 May 1998

ABSTRACT: The modeling of voltage-gated ion-channel proteins is a continuing challenge for force-field calculations because of the diverse range of interactions involved. In particular, current force fields are not parameterized for either ion–amino acid or amino acid–electric field interactions. To address the parameterization of ion–amino acid interactions, we have tested the use of empirical correction terms, derived from *ab initio* calculations of single amino acids (representing the peptide backbone) interacting with K^+ ions. Having demonstrated the utility of such an approach, we then extended the application to the amino acid side chains. The calculation of the interaction of K^+ with serine, cysteine, methionine, lysine, arginine, aspartate, histidine (uncharged), tyrosine, tryptophan, and phenylalanine, both completes the parameterization of the molecular environments contained in the amino acids, and allows specific comment on these ion–functional group interactions. The cation–π interactions were of particular interest, given recent proposals in the literature and the fear that force fields would not be able to treat such interactions. We present a comprehensive comparison of the *ab initio* (DFT [BLYP], 6-311 G**) and force field (CHARMm22.0) assessments of these interactions. J Comput Chem 19: 1515–1525, 1998

Keywords: cation–π; development; CHARMm22.0; correction term; force field

Correspondence to: D. F. Weaver; e-mail: weaver@chem.queensu.ca

Contract/grant sponsor: Natural Science and Engineering Research Council

CCC 0192-8651 / 98 / 131515-11

Introduction

Ion-channel proteins are macromolecules (> 20,000 atoms) that transverse cellular membranes, thus existing within a biological electric field. These dynamic molecules control the transmembrane movement of ions, thereby determining electrical activity and information flow in organs such as brain and heart. Although these proteins are fundamental to numerous disease states, theoretical structural analyses of their geometries and conformations (crucial to rational drug design) are hindered by the inability of current molecular modeling methods to treat such problems. Modeling transmembrane ion-channel proteins will require force fields that are parameterized for both ion–amino acid interactions and the presence of an electric field. Although there have been successful examples of incorporating electric fields into force-field-based molecular dynamics simulations, ion–amino acid interactions have proven more challenging. The problems encountered are twofold. First, the theoretically derived expressions to describe these interactions usually involve polarizability,[1–6] which increases computational cost tremendously and precludes the calculation of biological macromolecules. Second, once an equation is chosen, it is difficult to parameterize for cation–protein interactions due to the dependence of these interactions upon many atoms. The overall goal of this article is to devise a strategy for addressing the ion–protein interaction problem.

In devising a method to develop large-scale computational strategies for ion–protein interactions, various approaches are possible. Because *ab initio* (molecular orbital, density functional theory) methods simply cannot address problems of this magnitude, empirical (force field) methods must be pursued. Although it is possible to develop either an entirely new force field or to modify an existing force field, the latter approach is preferable and may be achieved by adding empirical correction terms to a current force field. Emprically derived correction terms (where ''term'' represents both the mathematical equation and the associated coefficients) have two major advantages. If the terms are selected from a series of simple functions (e.g., ranging from $1/d$ to $1/d^{12}$, where d is the distance between the specified atom and the ion) with a fixed coefficient for a given ion, they are computationally inexpensive and are well suited for the computational treatment of macromolecules. More importantly, the added terms describe a specific interaction by adding specific terms; hence, it is possible to avoid changing already defined parameters in the force field that were derived to reflect reasonable internal motions of proteins.

Accordingly, the specific goal of this article is to modify a current force field (CHARMm) by the addition of empirical correction terms of the form A/d_{ij}^{n} ($n = 1–12$). These empirical corrections are determined by fitting with multiple linear regression analyses to rigorous density functional theory calculations on model ion–amino acid interactions. In the current work we parameterize CHARMm to treat K^+ interactions with both peptide backbone and amino acid side chains. Potassium was selected because it is an important biological cation. This study is divided into four sections representing different molecular environments: part A discusses K^+ interactions with aliphatic, carbonyl, and amide NH functional groups of the peptide backbone; part B investigates the side chains of amino acids cysteine (Cys), serine (Ser), and methionine (Met); part C discusses the side chains of arginine (Arg), lysine (Lys), and aspartate (Asp); and part D investigates the side chains of phenylalanine (Phe), tyrosine (Tyr), tryptophan (Trp), and (uncharged) histidine (His).

The end result of this study is ''CHARMm—Channel,'' a version of CHARMm22.0 modified (via *ab initio* calculations) to treat K^+–peptide interactions by means of seven empirical correction terms. After applying the carbonyl (C and O) and aliphatic (C and H) backbone correction terms, only three additional terms are required for the sulfhydryl and hydroxyl functional groups; all other side chains are sufficiently well modeled by the existing force field. In particular, the cation–π interactions are well modeled (within our defined criteria) by Coulomb and van der Waals terms only.

Method

Ab initio (density functional theory) calculations were performed using Gaussian-94 (G94)[7] on an eight-node IBM parallel SP2 processor. CHARMm-22.0[8–10] and Quanta4.1[11] were installed on an IBM RS-6000/550 processor and were used for force-field calculations and molecule visualization, respectively. Correction terms were fitted with mul-

tiple linear regression analyses using Quattro Pro, version 6.

STEP PROCEDURE FOR DERIVING FORCE-FIELD CORRECTION TERMS

The derivation of the correction terms for the CHARMm force field followed a three-step procedure that was applied to the four peptide molecular environments designated previously.

Step I. Interactions between an ion and single amino acids were rigorously calculated throughout a three-dimensional grid using *ab initio* (density functional theory [DFT] with the BLYP[12–14] functional in G94) calculations. The three-dimensional grid was specified by systematically varying the ion around the functional groups of the amino acids. Thus, the ion was moved outward from the targeted functional group along the functional group axis in a radial (or ''R'') search. The distance between the functional group and the ion was then fixed and the ion was moved along an angular path from $-100°$ to $+100°$ at 5° increments in two directions, designated θ and ϕ (Fig. 1). The included graphs (Figs. 2–4) and tables refer to the ''R'' portion of the grid search only.

The 6-311G**[7] basis set from Gaussian-92 was used for amino acids (with the exception that the *p* coefficient on the *sp* orbitals of oxygen was 0.1 instead of 1.0). Diffuse functions were added to the oxygens of aspartate. The K^+ basis set was Ahlrichs' VDZ.[15] The DFT calculations were done with a fine grid. Initially, basis set superposition error (BSSE) was estimated with counterpoise correction.[16,17] The BSSE was found to be less than 1.5 kcal/mol with K^+. For the purposes of the present study (see ''Assumptions'') the error was considered small and counterpoise correction was not used.

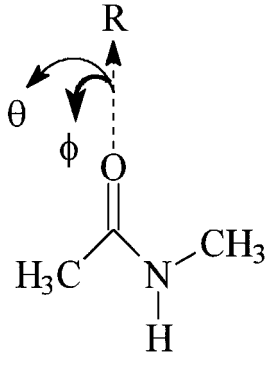

FIGURE 1. The conformation of *N*-methylacetamide with the R, θ, and ϕ grid search regions about the carbonyl functional group identified. These regions were similarly defined for all other functional groups and amino acids.

Step II. The same calculations were then performed with an empirical force field (CHARMm 22.0). In all calculations, the dielectric constant was set to one and no nonbonded cutoff was used. The all-atom representation was used throughout. The geometries and conformations of the amino acids used in both steps I and II were generated from the CHARMm-minimized default structures (with biological configuration) with —$COCH_3$ attached to the N terminal and with amidation of the C terminal.

Step III. Differences between the empirical and *ab initio* calculations were then quantified and used to derive empirical correction terms for the force-field equation. This was broken down into two tasks (A and B).

Task A. The difference between the force-field results and the *ab initio* results in the ''R'' and ''θ'' regions of the grid search was assessed. If the difference was greater than 5 kcal/mol, then it was fitted with a succession of functions ranging from $1/d$ to $1/d^{12}$ (where d is the distance between the ion and the specified atom type) using multiple linear regression (MLR) analysis. The best-fitting function (in terms of the square of the correlation coefficient and the observed curve shape) was then used as a correction term and the resulting behavior was checked against other regions of the grid search. This approach ensured that the most reproducible term was used and avoided ''overfitting'' (see ''Discussion''). Correction terms were not fitted to the points where the potassium ion approached within 2 Å of the amino acid because the CHARMm data was inordinately repulsive.

Task B. For side chain interactions, the previously determined backbone correction terms were applied, and the resulting differences from the *ab initio* results were determined. If these results were within 5 kcal/mol for uncorrected functional groups (i.e., amino acids Ser, Cys, Met, Lys, and Phe), then no further correction terms were derived; otherwise, further correction terms were derived to the CHARMm + backbone correction force field as outlined in Task A. The remaining aromatic amino acids were used to confirm the portability of results developed in Tasks A and B.

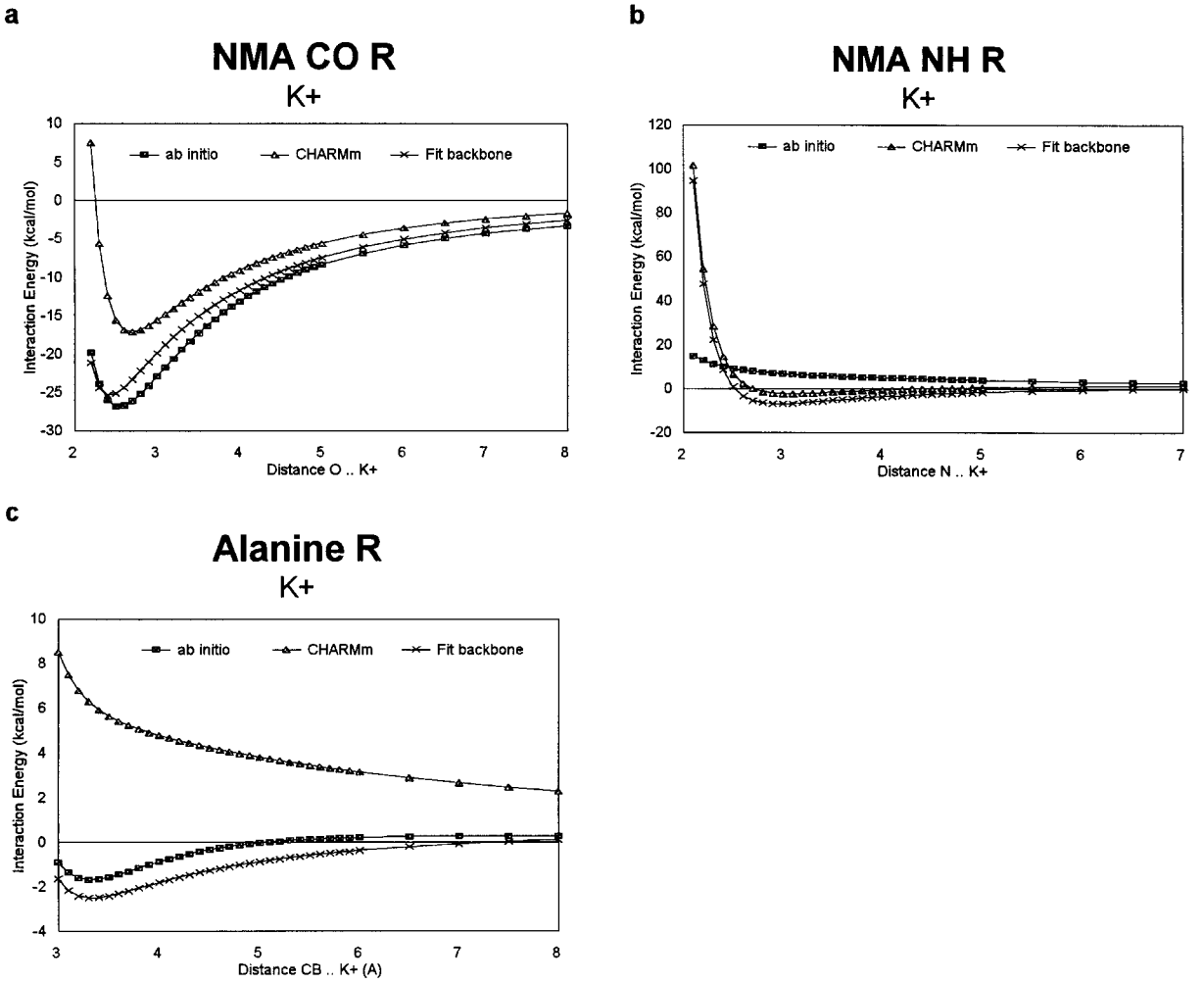

FIGURE 2. Interaction of K^+ with peptide backbone functional groups. Results are shown for *ab initio*, force field (CHARMm), and corrected force field (Fit backbone) calculations. (a) K^+ is incrementally moved farther away from O on the CO axis. (b) K^+ is incrementally moved farther away from the H on the NH axis. (c) K^+ is incrementally moved farther away from C_β on the $C_\alpha C_\beta$ axis.

This three-step procedure was applied sequentially to the four unique peptide environments defined in what follows.

Results

PEPTIDE BACKBONE CORRECTION TERMS—NMA AND ALA

The backbone correction terms were initially fitted to *ab initio* grid searches of *N*-methylacetamide (NMA) and alanine (Ala). NMA is a widely used peptide backbone mimic. The carbonyl grid search of NMA (Fig. 2a) provided the correct curve shape but incorrect magnitude. Thus, the minimum in the carbonyl radial direction was at the correct length, but was 15 kcal/mol less attractive then the *ab initio* result. The NH grid search (Fig. 2b) showed similar repulsive behavior when comparing the CHARMm and *ab initio* results. The Ala grid search provided parameters for the aliphatic carbons and hydrogens, which occur in both peptide backbones and side chains. Because aliphatic groups also occurred in NMA, the two systems were considered simultaneously. The $—CH_3$ radial grid search of Ala showed a large difference between the *ab initio* and CHARMm curve shapes. The *ab initio* search revealed a shallow minimum around 3.3 Å, whereas the CHARMm curve was entirely repulsive (Fig. 2c). Initial fitting indicated that a d^{-12} function between potassium ion and the carbonyl oxygen, and

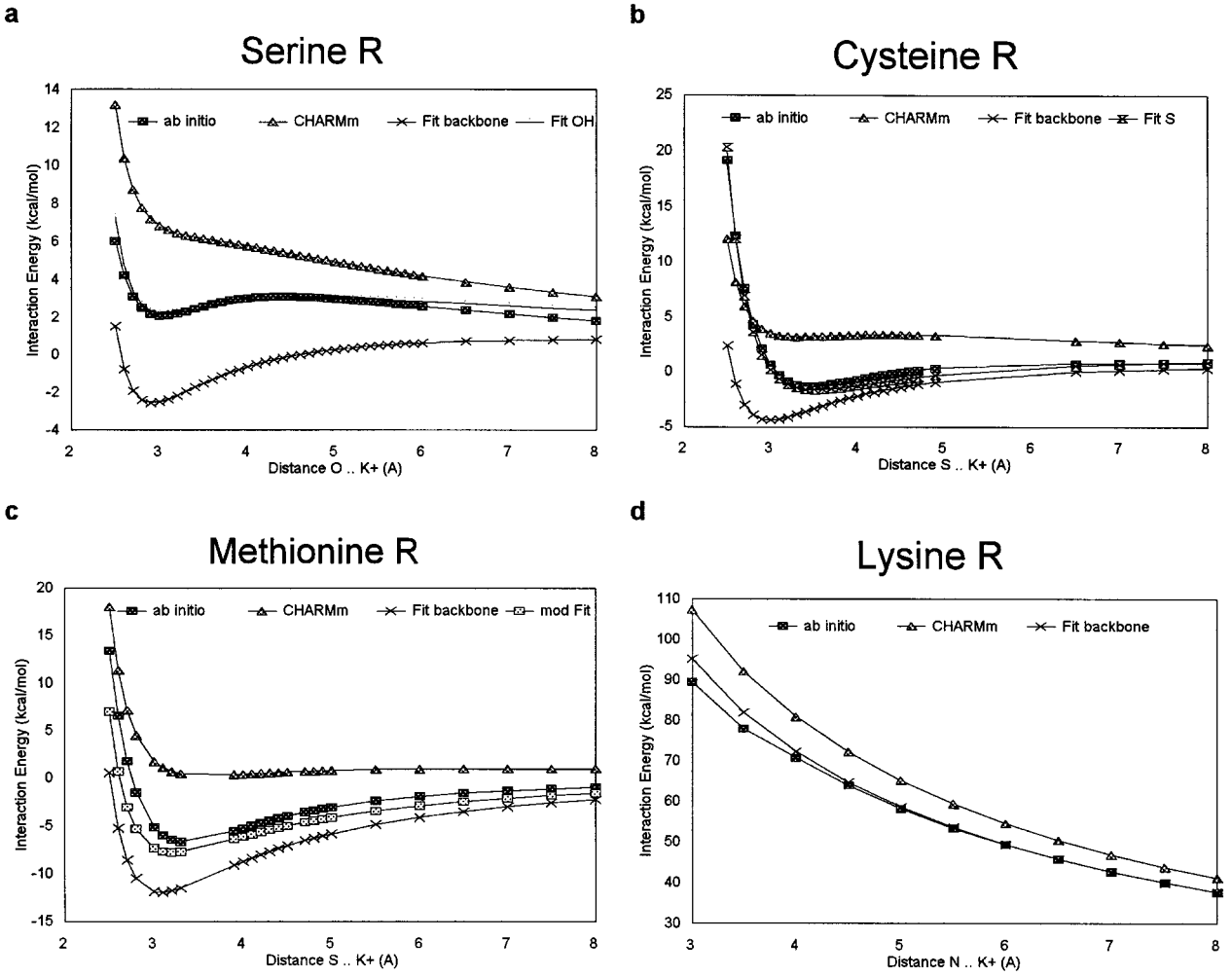

FIGURE 3. Interaction of K^+ with single amino acids. Results are shown for *ab initio*, force field (CHARMm), and corrected force field (Fit backbone) calculations. (a) K^+ is incrementally moved farther away from O on the $C_\beta O$ axis. The additional correction terms for the hydroxyl group are also shown (Fit OH). (b) K^+ is incrementally moved farther away from S on the SH axis. The additional correction term for the sulfur is also shown (Fit 8). (c) K^+ is incrementally moved farther away from S on the $C_\gamma S$ axis. Results are also shown for the backbone correction term not applied to the terminal methyl group (mod Fit). (d) K^+ is incrementally moved farther away from N on the CN axis.

potassium ion and the carbonyl carbon, was sufficient to fit the carbonyl difference. However, the —CH_3 grid search required terms that were d^{-2} for both the aliphatic carbon and hydrogen. Because the aliphatic correction term was so pervasive, it was necessary to determine the d^{-2} coefficients, and to then subtract this contribution from the carbonyl data before fitting the d^{-12} functions for the carbonyl carbon and oxygen. The resulting NH curve agreed within 5 kcal/mol and correction terms were not assigned (Fig. 2b). Thus, only the carbonyl and aliphatic corrections were implemented (see Table I for coefficients and Table II for the resulting correlation).

SERINE, CYSTEINE, AND METHIONINE

Serine (Ser), cysteine (Cys), and methionine (Met) all contain aliphatic side chains with either an OH or SR polar functional group. The Met sulfur is classified as a different atom type than the Cys sulfur, according to CHARMm and this convention was followed in the fitting procedure.

In the case of Ser, the CHARMm curve was almost entirely repulsive in the radial direction, whereas the *ab initio* curve showed a distinct minimum (Fig. 3a, Table II). The backbone-corrected CHARMm curve also revealed a minimum, but was too attractive. A d^{-1} correction term for both

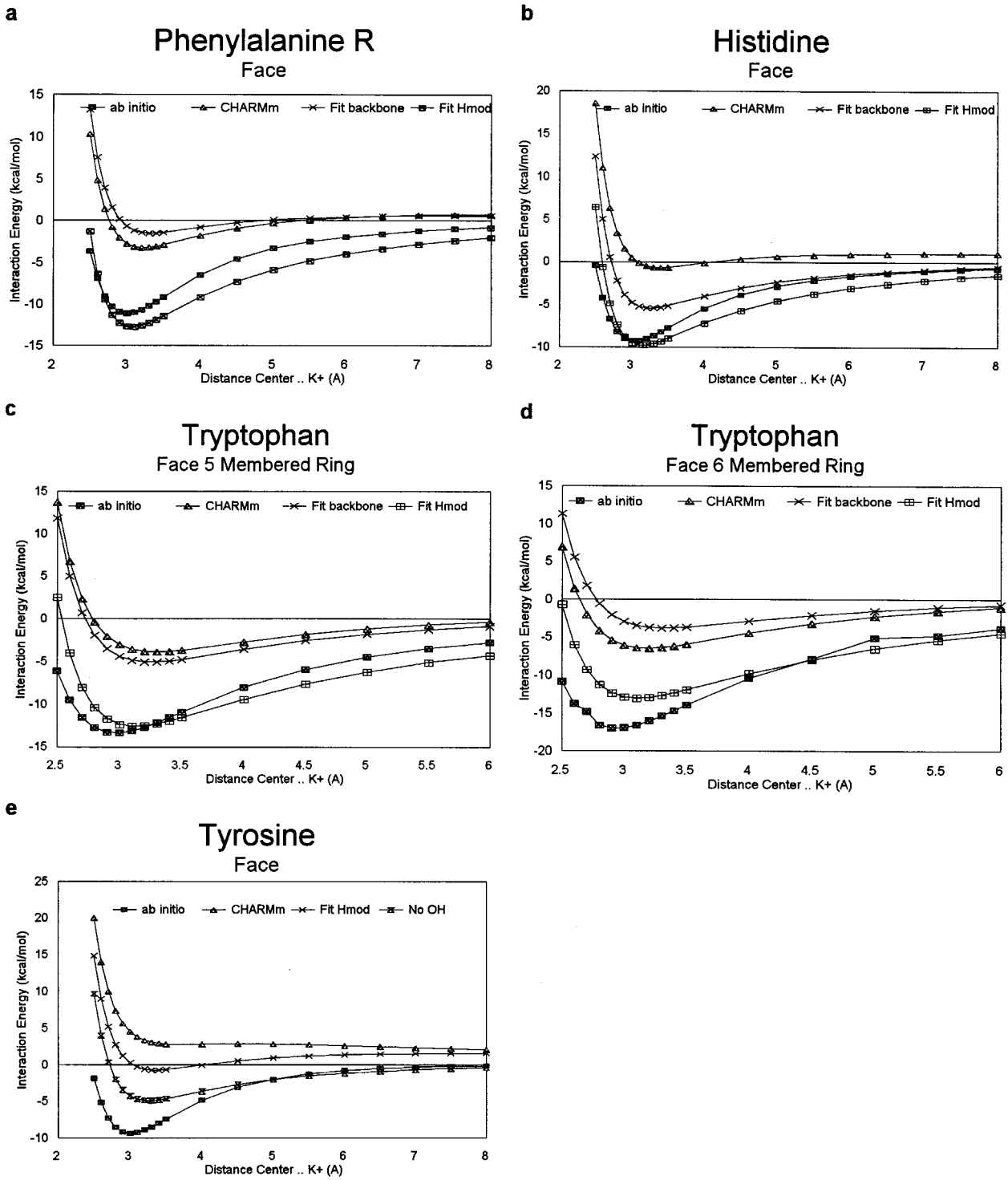

FIGURE 4. Interaction of K^+ with single amino acids incorporating π electrons. Graphs (a)–(e) use K^+. Results are shown for *ab initio*, force field (CHARMm), corrected force field (Fit backbone), and corrected force field with distinction of the ''aromatic hydrogens'' calculations (Fit Hmod). (a) K^+ is incrementally moved farther above the center of the plane of the ring. (c) K^+ is incrementally moved farther away from N above the plane of the ring. (c) K^+ is incrementally moved farther above the center of the plane of the five-membered ring. (d) K^+ is incrementally moved farther above the center of the plane of the six-membered ring. (e) K^+ is incrementally moved farther above the center of the plane of the ring. Also included is the curve that does not include the hydroxyl correction terms derived for serine (No OH).

TABLE I.
Correction Terms Derived for CHARMm22.0 Force Field.

CHARMm atom type (code #)	Functional exponent	Coefficient: potassium
Carbonyl carbon [14]	−12	68395.7
Carbonyl oxygen [40]	−12	−303054.9
Aliphatic carbon [10][a]	−2	−143.95
Aliphatic hydrogen [3][a]	−2	33.07
Hydroxyl oxygen [45][b]	−1	−71.61
Hydroxyl hydrogen [8][b]	−1	−51.73
Cysteine sulfur [70]	−8	26475.9

[a] Does not include the terminal CH_3 group on methionine or hydrogens bonded to ring carbons.
[b] Applies to serine only.

the hydroxyl oxygen and hydrogen was applied and found to provide very good agreement (Tables I and II).

The Cys results were similar to those for Ser. In this case, CHARMm again failed to show a minimum in the radial direction, whereas the *ab initio* data did (Fig. 3b). The backbone-corrected CHARMm curve showed a minimum that was too attractive. A d^{-8} correction term for sulfur agreed with the *ab initio* result (Tables I and II).

Finally, methionine also showed a distinct minimum in the *ab initio* radial calculations, whereas the CHARMm result did not (Fig. 3c). The backbone-corrected CHARMm data also showed a more attractive minimum. Functions on the methionine sulfur did not offer any improvement. The agreement was improved when the aliphatic correction term was not applied to the terminal CH_3 on sulfur (Fig. 3c). The application of the correction term to the α and β carbons and hydrogens only was also examined and was not found to be more advantageous (data not shown). By removing the terminal methyl from the backbone correction, no additional correction terms were required (Table II).

Overall, Cys and Ser required additional correction terms for the side chain functional groups (Table I). In both cases the CHARMm curve is repulsive, whereas the *ab initio* results indicate distinct minima. Applying a correction term specific for the Met sulfur was not necessary, and simply not applying the aliphatic correction term to the terminal methyl of the Met side chain was sufficient to provide good agreement.

TABLE II.
Square of Correlation Coefficients between CHARMm and Modified CHARMm Predictions and *Ab Initio* Calculations.

Amino acid	Adjusted	Original
NMA CO	0.976	0.237
NMA NH	0.992	0.974
Ala	0.958	0.736
Ser	0.902	0.483
Cys	0.996	0.929
Met	0.908	0.869
Lys	0.998	0.998
Arg	0.999	0.999
Asp	0.999	0.999
Phe side	0.914	0.883
Phe face	0.961	0.241
Trp five-ring	0.556	0.111
Trp six-ring	0.537	0.205
His face	0.747	0.101
His side	0.284	0.201
Tyr OH	0.999	0.229
Tyr face	0.924	0.007

LYSINE, ARGININE, AND ASPARTATE

The charged groups, lysine (Lys), arginine (Arg), and aspartate (Asp), interactions with K^+ should be dominated by charge–charge interactions. For brevity, only the representative lysine graph is shown (Fig. 3d).

The CHARMm interaction energy curve for K^+ and lysine mimics the *ab initio* curve shape quite consistently. The backbone-corrected CHARMm force field overlaps the *ab initio* curve almost exactly. No additional functions were required. The arginine results were very similar to those for lysine. Again, the backbone-corrected CHARMm force field agreed with the *ab initio* curve and no additional functions were required (Table II).

Similar to the positively charged amino acids, no additional functions were required for aspartate. The backbone correction was sufficient to give good agreement with the *ab initio* curve (Table II).

PHENYLALANINE, TRYPTOPHAN, HISTIDINE, AND TYROSINE

The cyclic unsaturated side chains are capable of interacting with ions through the electron density in their π orbitals, hence giving rise to "cation–π" interactions.

The phenylalanine–K^+ grid search above the face of the ring indicated a distinct minimum of approximately −12.5 kcal/mol at 3-Å separation. The CHARMm-generated data indicated a very

shallow −2.5 kcal/mol minimum at 3.5 Å (Fig. 4a). The backbone correction did not result in significant improvement. However, the CHARMm atom classification considers hydrogens bonded to aliphatic carbons and hydrogens bonded to aromatic carbons to be identical (i.e., both have a code number of 3). By distinguishing these ''aromatic'' hydrogens so that the backbone correction term was not applied to them (i.e., by reassigning the code number), the resulting curve agreed well with the *ab initio*-generated curve and no additional correction terms were required (Fig. 4a). In addition, the modified correction term was also checked against searches along the side of the phenylalanine ring and found to yield reasonable results (Table II).

The tryptophan results were similar. After distinguishing the ''aromatic'' hydrogens, the *ab initio* face searches above the five- and six-membered rings agreed well with the modified correction term and no additional term was required (Fig. 4d and e, Table II). *Ab initio* data in the NH radial direction of the five-membered ring and the lone CH on the five-membered ring also yielded good agreement (data not shown) with the corrected CHARMm force field.

Histidine (uncharged) was studied above the face of the ring and along the N direction (N without a hydrogen bonded to it). Again after distinguishing the hydrogens bonded to the ring carbons, the backbone-corrected CHARMm data agreed quite well with the *ab initio* results (Fig. 4b). No additional functions were required (Table II).

Tyrosine was grid searched above the face of the ring and in a radial direction from the hydroxyl group. The face search yielded a minimum of ~ -10 kcal/mol by *ab initio* calculation. The CHARMm result had no minimum. The correction term including the hydroxyl from serine was applied (aromatic hydrogens were again distinguished). However, better agreement was obtained without the OH functions (Fig. 4e, Table II). Simply applying a single oxygen or hydrogen function fitted to serine did not remedy the problem, indicating that the error was not due to overfitting of the serine hydroxyl grid search (see ''Discussion'').

Thus, *none* of the unsaturated side chains required additional correction terms. Inclusion of the backbone correction and distinguishing ''aromatic'' hydrogens yielded acceptable results (Table II, Fig. 4).

SUMMARY OF CORRECTION TERMS

In summary, only seven atom types required correction terms (Table I). Considering the number of diverse environments that have been tested in these studies, this was unexpected. The form of the correction terms is as follows

$$E = A/d_{ij}^{n} \quad (1)$$

where d_{ij} = distance between the specified ion (i) and the specified CHARMm atom type (j).

Analyses of the square of the correlation coefficient between the various fitted curves with the *ab initio* results as compared to the r^2 between the original CHARMm curves and the *ab initio* results are included in Table II. The r^2 values should be considered in conjunction with the included graphs (Figs. 2–4).

Discussion

The interaction of a potassium ion with a variety of amino acids encompasses a diverse range of molecular environments. Interpretation of the results in each environment requires an understanding of the assumptions and philosophy utilized in the study (''Assumptions''). Furthermore, each environment yielded its own conclusion(s). Consideration of the peptide backbone, charged, and uncharged saturated side chains, respectively (see following subsections), provides comment on CHARMm parameterization of these environments. Finally, the unsaturated cyclic side chains, which arguably provide the most complex environment, must also be addressed in light of recent studies of cation−π interactions (see later subsection).

ASSUMPTIONS

A number of implicit assumptions are made when computational methods are utilized. The degree of accuracy used in the original force field parameterization and the conformational dependence of the *ab initio* calculations both limit the fitting criteria to within 5 kcal/mol of the correct curve shape. Furthermore, attempts to fit the *ab initio* data very precisely by adding more terms resulted in overfitting. That is, the errors in the coefficients were larger than the coefficients themselves, and the results were not portable outside of the fitted region. Thus, in the interests of achieving

a robust and portable force-field modification, more general fitting criteria were applied.

PEPTIDE BACKBONE CORRECTION TERMS

The largest and most surprising correction was required for the aliphatic group of alanine. This correction term is likely to be due to electrostatic interactions because the functional form of the correction is d^{-2}. The carbonyl correction was much smaller, and likely related to a van der Waals type of interaction (because the functional form was d^{-12}). The amide NH function did not require any correction. The validity of these correction terms is evident from the relatively few additional terms required by side chain molecular environments. In particular, the success with the straight chain charged amino acids, as discussed below, demonstrates the validity of the aliphatic and carbonyl correction terms.

SERINE, CYSTEINE, AND METHIONINE

All three of the amino acids with uncharged linear side chains were not well modeled by the original CHARMm force field (no minima were found), whereas the backbone-corrected CHARMm curves were overly attractive. The hydroxyl group in Ser required the most pervasive (d^{-1}) correction, whereas the Cys was best modeled with a d^{-8} function on the sulfur. However, Met could not be ''adjusted'' with a sulfur function, but rather required removing the terminal methyl from the backbone correction. Thus, the corrections can be ranked in the order Ser > Cys > Met. However, in terms of curve shape (Fig. 3a–c), Cys required the most correction. It is of interest to note that the polarizability increases as OH < SR < SH. Thus, the most pervasive corrections (i.e., OH) are more likely related to deficiencies in van der Waals (VDW) and Coulomb parameters of the CHARMm force field. However, the different curve shape in the Cys case may be related to polarizability.

LYSINE, ARGININE, AND ASPARTATE

The charged amino acids showed excellent agreement with the *ab initio* results (Fig. 3d). This should be expected because the charge–charge electrostatics should dominate the interaction, and hence the force field should be able to model this interaction very well with the Coulomb term. It is of interest to note that the (aliphatic) backbone correction was beneficial in all cases, supporting the parameters and functions chosen previously.

PHENYLALANINE, TRYPTOPHAN, HISTIDINE, AND TYROSINE: CATION–π INTERACTIONS

The final molecular environment that must be addressed concerns cation–π interactions. Cation–π interactions are the subject of intense interest in biological systems at present,[4,5] and many calculations have been done on model systems of benzene, phenol, and indole interacting with various ions.[18,19] Similarly, the electrostatic characteristics of benzene have been well studied both experimentally and computationally.[20,21] However, the origin of the cation–π interaction is a point of debate. Dougherty and associates ascribe the variation in cation–π binding energy with predominantly electrostatic effects ($\sim$ 60%).[5,22,23] In this case, electrostatic effects are defined as the interaction of the aromatic group's partial charges, dipoles, quadrupoles, etc., with positive charge on the ion.[3] It does not include polarization effects that involve the *response* of the aromatic groups to the positive charge and can result in induced dipole–ion interactions as well as many others.[3] Kollman has also had success in fitting benzene–cation interactions very accurately (at the expense of CPU time) with three-body term force fields including by polarizability.[24–26] It has been argued that, if the system can be well modeled, then the underlying classical effect must occur and thus polarization is proposed to play a role. Dougherty has stated that there is a constant 12-kcal/mol interaction in the systems studied, which could be ascribed to polarization effects,[23] whereas the variation in the data is due to electrostatic effects.

The present study considers the relevant amino acids with a flexible basis set (known to be required to account for polarization effects), and some degree of correlation (using DFT). The interaction energy between K^+ and the aromatic amino acids are ranked in the order tyrosine (−10 kcal/mol) < phenylalanine (−12.5 kcal/mol) < tryptophan (−13-kcal/mol five-membered ring, −17-kcal/mol six-membered ring). The model systems are ranked with benzene (−19.2 kcal/mol) < phenol < indole according to their electrostatic potentials using 6-311 + G** Hartree–Fock calculations.[5] These differences may be due to the influence of the backbone in the amino acids. From this study we cannot assign the origin of the interaction of K^+ with the aromatic amino acids. However, whatever the interaction is, it is well mod-

eled by CHARMm22.0 to within 5 kcal/mol once the backbone correction is included (within 10 kcal/mol for Tyr). This may suggest that the interaction is indeed primarily electrostatic because the distribution of partial charges could be expected to model dipoles and quadrupoles fairly well. However, polarization effects could also be present and are either well modeled by the current functions or contribute less than 5 kcal/mol. Most importantly, however, this effect is not as poorly modeled by current force fields as was initially expected.

This empirically discovered result regarding the sufficiency of force fields with ''aromatic'' hydrogens has been previously noted in the literature.[27] Although the interaction studied concerned ammonium cations, vs. Na^+ or K^+, it also illustrated that the charges on the ''aromatic'' hydrogens were required to model the cation$-\pi$ interaction as compared to a united-atom approach. The cited work, however, could not comment on the degree of sufficiency of the charges assigned to carbons and hydrogens of the aromatic group (comparison was made between experimental and calculated binding energies). Because our comparison involved *ab initio*-derived interaction energies, it is possible to comment on the degree of agreement —in this case generally within 5 kcal/mol.

Finally, tyrosine requires individual comment. The agreement with tyrosine is not as good as with phenylalanine and tryptophan (Fig. 3a–e). Moreover, the hydroxyl grid search in tyrosine is quite different from that for serine; however, it is well modeled as long as the OH correction term, derived for serine, is not included. These differences in behavior suggest that there is a fair degree of interaction between the hydroxyl and aromatic group.

Thus, the expected deficiency of the force field in modeling the cation$-\pi$ interaction is not found to occur, within the fitting criteria and with the backbone correction.

Conclusion

To perform detailed studies of voltage-gated ion channels, force fields that accurately treat ion–amino acid interactions are required. To this end, a three-step procedure (I–III) was developed that allowed the empirical determination of correction terms for the ion–amino acid interactions of four peptide environments (parts a–d).

As a result of this approach, it is possible to observe both the deficiency and sufficiency of the force field. We determined that the force field incorrectly assessed the degree of interaction between aliphatic groups and K^+ and, to a lesser extent, carbonyl groups and K^+. However, the interaction of charged groups with K^+ was very well modeled. Additional correction terms were required for the sulfhydryl and hydroxyl groups of cysteine and serine. Most surprisingly, the combination of van der Waals and Coulomb forces, as prescribed by the CHARMm parameters, were sufficient to model the cation$-\pi$ interaction to within 5 kcal/mol. The correction terms applied (Table I) were empirically derived, and allowed the separation of the backbone contribution from that of the cation$-\pi$ interaction.

It is important to note that any two sets of data can always be empirically fit. Thus, it is necessary to ask whether the resulting correction terms are meaningful. This can be assessed by considering the portability of the correction terms. If the corrections were arbitrary, they would not be applicable to other amino acid environments, such as side chains, nor to the other regions of the grid search. In all cases, correction terms were only accepted if, after being fit to the ''R'' and positive ''θ'' regions of the grid search, they were portable to the negative θ and ϕ regions. Furthermore, the extension of the backbone correction terms to the amino acid side chains illustrates that the correction is for inherent structure in the data. This issue has been addressed further in the extension of the backbone correction terms to the interaction of glycylglycine with K^+, and to the interaction of Na^+ with NMA, Ala, and glycylglycine. The interaction of K^+ with glycylglycine was examined with the previously derived correction terms, and the corrected force field successfully mimicked the *ab initio* result. Similarly, the extension of the correction terms to Na^+ was tested by rederiving the ion-specific coefficient, but using the same exponential form as was found for K^+; the resulting corrected force field and *ab initio* results agreed well. Consequently, the two sets of data (i.e., the force field and *ab initio* results) have not been arbitrarily fitted and the correction terms represent inherent interactions between the ions and the amino acid functional groups.

In conclusion, CHARMm-Channel adjusts the CHARMm22.0 force field in a manner that allows the interaction of a potassium ion with the potassium ion channel to be examined in a reliable manner. Correction terms were required for aliphatic carbons and hydrogens, carbonyl carbons

and oxygens, the serine hydroxyl group, and the sulfur of cysteine, whereas unsaturated side chains were found to be surprisingly well modeled by CHARMm. The strategies used to develop this particular parameterization can also be applied to other biologically relevant cations (e.g., Na^+) or to other force fields (e.g., AMBER, MM2). The conclusion of this empirical parameterization also allows us to begin to simulate atomic level detail models of ion channels and ion channel permeation.

Acknowledgments

O.D. acknowledges a Natural Science and Engineering Research Council (NSERC) scholarship. D.F.W. acknowledges an Ontario Ministry of Health Career Scientist Award. We thank the referee #1 for bringing ref. 27 to our attention.

References

1. B. Roux and M. Karplus, *J. Comput. Chem.*, **16**, 690 (1995).
2. C. E. Dykstra, *Chem. Rev.*, **93**, 2339 (1993).
3. G. Naray-Szabo and G. G. Ferenczy, *Chem. Rev.*, **95**, 829 (1995).
4. R. A. Kumpf and D. A. Dougherty, *Science*, **261**, 1708 (1993).
5. D. A. Dougherty, *Science*, **271**, 163 (1996).
6. W. Sokalski, *Amino Acids*, **7**, 19 (1994).
7. M. F. Frisch, G. W. Trucks, H. B. Schlegel, et al., *Gaussian-94, rev. B*, Gaussian, Inc., Pittsburgh, PA, 1995.
8. *CHARMm22.0*, MSI Simulations, Boston, MA, 1993.
9. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, et al., *J. Comput. Chem.*, **4**, 187 (1983).
10. F. A. Momany and R. Rone, *J. Comput. Chem.*, **13**, 888 (1992).
11. *Quanta4.1*, MSI Simulations, Boston, MA, 1994.
12. C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B.*, **37**, 785 (1988).
13. A. D. Becke, *Phys. Rev. A*, **38**, 3098 (1988).
14. B. Miehlich, A. Savin, and H. Stoll, and H. Preuss, *Chem. Phys. Lett.*, **157**, 200 (1989).
15. Ahlrichs VDZ (A. J. H. Wachters, *J. Chem. Phys.*, **52**, 1033 [1970]). Downloaded from http://www.emsl.pnl.gov:2080/cgi-bin/run-bsform-post.
16. M. J. Frisch, J. E. Del Bene, J. S. Brinkley, and H. F. Schaefer III, *J. Chem. Phys.*, **84**, 2279 (1986).
17. D. W. Schwenke and D. G. Truhlar, *J. Chem. Phys.*, **82**, 2418 (1985).
18. J. Y. Lee, S. J. Lee, H. S. Choi, S. J. Cho, K. S. Kim, and T. Y. Ha, *Chem. Phys. Lett.*, **232**, 67 (1995).
19. K. S. Kim, J. Y. Lee, S. J. Lee, T. K. Ha, and D. H. Kim, *J. Am. Chem. Soc.*, **116**, 7399 (1994).
20. G. R. Dennis and G. L. D. Ritchie, *J. Phys. Chem.*, **95**, 656 (1991).
21. I. R. Gentle and G. L. D. Ritchie, *J. Phys. Chem.*, **93**, 7740 (1989).
22. P. S. Kearney, L. S. Mizoue, R. A. Kumpf, J. E. Forman, A. McCurdy, and D. A. Dougherty, *J. Am. Chem. Soc.*, **115**, 9907 (1993).
23. S. Mecozzi, A. P. West Jr., and D. Dougherty, *J. Am. Chem. Soc.*, **118**, 2307 (1996).
24. J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 4177 (1995).
25. L. X. Dang, J. E. Rice, J. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **113**, 2481 (1991).
26. J. Caldwell, L. X. Dang, and P. A. Kollman, *J. Am. Chem. Soc.*, **112**, 9144 (1990).
27. P. H. Axelsen, *Isr. J. Chem.*, **34**, 159 (1994).